

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 8月27日

出 願 番 号

Application Number:

平成11年特許願第241024号

出 願 人

Applicant(s):

株式会社日立製作所

USSN 09/642,817

MATTINGLY, STANGER + MALUR

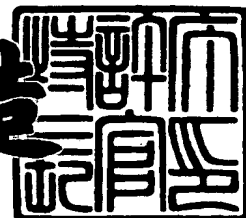
(703) 684-1120

DKT: ASA-911

2000年 8月25日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3067679

【書類名】 特許願

【整理番号】 K99009461

【提出日】 平成11年 8月27日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 12/00

【請求項の数】 2

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 北村 学

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 山神 憲司

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 村上 達也

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 計算機システム

【特許請求の範囲】

【請求項 1】

複数の計算機と、前記複数の計算機に接続される記憶装置サブシステムとを有し、前記記憶装置サブシステムは 1 または複数のデバイスを有し、前記複数の計算機のうち 1 つの計算機は前記複数の記憶装置サブシステムの有する前記複数のデバイスの情報を管理する管理手段を有し、前記複数の計算機は、前記管理手段に対しデバイスの割り当てを要求する要求手段を有する計算機システムであって、前記管理手段は、前記デバイスの割り当てを要求する手段からの要求に基づき前記複数の記憶装置サブシステム内の複数のデバイスの中から適切なデバイスを選択する選択手段を有し、前記複数の計算機は、前記選択手段により選択された前記デバイスをアクセスできるよう当該計算機の設定を変更する変更手段を有することを特徴とする計算機システム。

【請求項 2】

複数の計算機と、前記複数の計算機に接続される記憶装置サブシステムとを有し、前記記憶装置サブシステムは 1 または複数のデバイスと、該複数のデバイスの情報を管理する管理手段を有し、前記複数の計算機は、前記管理手段に対しデバイスの割り当てを要求する要求手段を有する計算機システムであって、前記管理手段は、前記要求手段からの要求に基づき前記複数の記憶装置サブシステム内の複数のデバイスの中から適切なデバイスを選択する選択手段を有し、前記複数の計算機は、前記選択手段の選択した前記デバイスをアクセスできるよう当該計算機の設定を変更する変更手段を有することを特徴とする計算機システム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、情報処理システムなどにおける記憶装置システムのディスク共有方法に係り、特に、ファイバチャネルインタフェースを有する記憶装置サブシステムで構成された計算機システムにおいて、デバイスの割り付けを自動化する方法

に関する。

【0002】

【従来の技術】

近年、企業内で扱われる情報量は飛躍的に増大し、それに伴いデータを記憶するディスク装置などの容量も増加の一途をたどっている。そのため数TB（テラバイト）の容量を持つディスク装置も珍しくなくなっている。例えば、特開平9-274544号公報には、1つの記憶装置サブシステムをさまざまな種類のデバイスで構成する技術が開示されている。ここに開示される技術は、具体的には、ディスク装置に関するもので、ホストからアクセスされるデバイス（論理ディスク）として、RAID（Redundant Arrays of Inexpensive Disks）におけるRAID5のものと、RAID1のものを混在させ、あるいは、論理ディスクを構成するための実際の磁気ディスク（物理ディスク）として、低速なものと高速なものを混在させて構成される。ユーザは、各デバイスのアクセス頻度を見て、ドライブの種類を使い分けることができる。

【0003】

また、ファイバチャネル技術が現れたことにより、複数のホスト計算機、複数の記憶装置を1つのファイバチャネルケーブルで接続した構成をとる計算機システムも現れてきた。こういった計算機システムでは、それぞれのホストは、ファイバチャネル上の任意の記憶装置に直接アクセスできる。このため、従来の各ホストがそれぞれに記憶装置を持つ構成に比べ、データの共有やネットワーク負荷の低減が期待できる。

【0004】

【発明が解決しようとする課題】

上述した従来技術では、各ホストがアクセスできるデバイスの数や種類を飛躍的に増加させることができる。しかし、ホストがアクセスできるデバイスの数や種類が増加するに伴い、各ホストでのデバイスの管理が困難になってくる。1つのホストから多くのデバイスにアクセスできる利点の一方で、ユーザにとっては、ある業務でどのデバイスを使用すれば良いかを選択することが困難となる。特に、ファイバチャネルで接続された計算機システムの場合、自ホストが本来使用

していないデバイスまでもアクセスできることとなり、間違えると、他のホストで使用しているデバイスに不正アクセスを起こし、データを破壊してしまうこともあり得る。特開平 10-333839 号公報では、このような問題を解決するため、ファイバチャネルで接続された記憶装置が特定ホストからのアクセスだけを許可する方法を開示している。しかし、記憶装置、デバイスが複数になった場合や、異なる種類のデバイスが混在する場合には、処理が複雑で有ることには変わりはなく、各ホストがそれぞれデバイスの種類などを常に意識する必要がある。

【0005】

本発明の目的は、各ホストが用途に合ったデバイスを必要な時に必要な分だけ利用できるように、デバイスの設定、デバイスの各ホストへの割り付けを容易に行えるようにすることにある。

【0006】

【課題を解決するための手段】

本発明における計算機システムは、その好ましい態様において、複数の計算機と、これら複数の計算機に接続する記憶装置サブシステムとを有する。記憶装置サブシステムは、複数のディスク装置（デバイス）と複数のインタフェースとを有し、各計算機に接続される。複数の計算機の 1 つは、記憶装置サブシステム内のデバイスと、各計算機と記憶装置サブシステムの接続関係の情報を保持する管理手段を有する。各計算機は、新たにデバイスが必要なとき、管理手段にその容量や種類を通知する。管理手段は、それを受けて、要求に合ったデバイスを選択する。そして選択したデバイスが計算機からアクセスできるように、記憶装置サブシステムに対し所定の情報を設定するように指示する。その後、デバイス割り当ての要求のあった計算機に所定の情報を返却し、要求のあった計算機で、はその情報に基づき、計算機の設定を変更し、当該計算機で割り当てられたデバイスが使用できる様にする。

【0007】

また、複数のホストコンピュータと複数の記憶装置サブシステムがファイバチャネルで接続された構成では、管理手段は、記憶装置サブシステム内のデバイス

の情報を有する。各記憶装置サブシステムは管理手段から指定された計算機に対してアクセスを許可する手段を有する。各計算機は、新たにデバイスが必要なとき、管理手段にその容量や種類を通知する。管理手段は、要求に合ったデバイスを選択し、選択したデバイスが計算機からアクセスできるように、記憶装置サブシステムに対し当該計算機からのアクセスを許可するよう指示する。その後、デバイス割り当ての要求のあった計算機に所定の情報を返却する。要求のあった計算機ではその情報に基づき、計算機の設定を変更し、計算機で割り当てられたデバイスが使用できる様にする。

【0008】

【発明の実施の形態】

ー第1の実施形態ー

図1は、本発明を適用した計算機システムの一実施形態における構成例を示すブロック図である。計算機システムは、複数のホスト計算機、ホスト計算機1a、ホスト計算機1b、…ホスト計算機1n（総称してホスト1と呼ぶ）と、ホスト1に接続される記憶装置サブシステム2、管理用ホスト計算機3、ネットワーク4、及び遠隔地に配置される記憶装置である副記憶装置5を有して構成される。

【0009】

ホスト計算機1a、1b、…は、CPU、メモリなどを有する計算機であり、メモリに格納されたオペレーティングシステム、アプリケーションプログラムをCPUが読み出して実行することで、所定の機能を達成する。

【0010】

記憶装置サブシステム2は、複数のディスク21、ディスクコントローラ22、ホスト計算機1と接続する複数のポート23、副記憶装置5と接続するためのインタフェース24、ネットワーク4と接続するネットワークインタフェース25を有している。本実施例における記憶装置サブシステム2は、複数のディスク21をまとめて1または複数の論理デバイスとしてホスト計算機1に見せかける。もちろん、個々のディスク21を1つの論理デバイスとしてホスト計算機1に見せるようにしても良い。

【0011】

ポート23としては、たとえば接続されるホスト計算機1が、いわゆるオープンシステムの計算機であればSCSI (Small Computer System Interface) などのインタフェースが用いられる。一方、ホスト計算機1が、いわゆるメインフレームであれば、ESCON (Enterprise System CONnection) などのチャンネルインタフェースが用いられる。それぞれのポート23は、同一のインタフェースであっても異なるものが混在していてもかまわない。本実施例では、全てのポート23にインタフェースとして、SCSIを用いるものとして説明する。

【0012】

ディスクコントローラ22は、プロセッサ221、キャッシュメモリ222、制御メモリ223を有している。プロセッサ221は、ホスト計算機1からのアクセスやディスク21の制御、特に、記憶装置サブシステム2がホスト計算機1に対してディスク21単体ではなく、ディスクアレイの様に複数のディスクをまとめて1または複数の論理デバイスに見せかけている場合には、その処理や管理などを行う。また、ディスクコントローラ22は、ネットワークインタフェース25を介して管理用ホスト計算機3との通信を行う。

【0013】

キャッシュメモリ222は、ホスト計算機1からのアクセス処理速度を高めるため、頻繁に読み出されるデータを格納したり、あるいはホスト計算機1からのライトデータを一時的に格納したりする。また、キャッシュメモリ222の一部を、1または複数の論理的なディスクに見せかけ、磁気ディスクへのアクセスが不要なデバイスを定義し、利用することもできる。

【0014】

制御メモリ223は、プロセッサ221が実行するプログラムを格納したり、ディスク21やそれを組み合わせてできる論理デバイスの管理を行うための情報を格納しておくために使用される。

【0015】

各ホスト計算機1a、1b…には、ボリュームマネージャ11と呼ばれるソフト(プログラム)が配置される。ボリュームマネージャ11は、管理用ホスト計

算機 3 に配置される管理マネージャ 31 と通信しあって動作する。また、各ホスト計算機 1 は、インタフェース (I/F) 12 を持っており、インタフェース 12 により記憶装置サブシステム 2 のポート 23 と接続する。

【0016】

次に、記憶装置サブシステム 2 内の論理デバイスの管理形態について説明する。

【0017】

前述の通り、記憶装置サブシステム 2 では、複数のディスク 21 をまとめて 1 または複数の論理デバイスとして、あるいは、個々のディスク 21 を 1 つの論理デバイスとしてホスト計算機 1 に見せかける。また、キャッシュメモリ 222 の一部の領域を 1 または複数の論理デバイスとしてホスト計算機 1 に見せる。記憶装置サブシステム 2 内のディスク 21 の数と論理デバイスの数に相関関係はない。

【0018】

記憶装置サブシステム 2 内では、図 2 のような論理デバイス管理テーブルを持って論理デバイスを管理している。論理デバイス管理テーブルでは、論理デバイス番号 61 に対し、サイズ 62、構成 63、状態 64、バス 65、ターゲット ID 66、LUN 67 の項目の組が管理される。サイズ 62 には、当該論理デバイスの容量を表す情報が設定される。構成 63 には、当該論理デバイスの構成を示す情報、たとえばディスク 21 が RAID (Redundant Arrays of Inexpensive Disks) を構成し、それが論理デバイスに割り当てられている場合、RAID1、RAID5 などの種別を示す情報が設定される。その他、キャッシュメモリを論理デバイスとして割り当ててあれば「キャッシュ」、単体のディスクを割り当てている場合には「単体ディスク」を示す情報が構成 63 に設定される。状態 64 には、当該論理デバイスの状態を示す情報が設定される。状態としては、「オンライン」、「オフライン」、「未実装」、「障害オフライン」が存在する。「オンライン」は、当該論理デバイスが正常に稼動し、ホストからのアクセスが可能な状態であることを示す。「オフライン」は、当該論理デバイスは定義され、正常に稼動しているが、ホストからのアクセスはできない状態にあることを示す。

。この状態は、以前ホスト計算機1で使用されていたが、ホスト計算機1でそのデバイスが不要となって使われなくなった場合に相当する。「未実装」は、当該論理デバイスが定義されておらずホストからのアクセスはできない状態にあることを示す。また「障害オフライン」は当該論理デバイスに障害が発生してホストからのアクセスができないことを示す。

【0019】

パス65には、当該論理デバイスが複数のポート23のどのポートに接続されているかをあらわす情報が設定される。各ポート23には、記憶装置サブシステム2内で一意な番号が割り振られており、「パス」欄には論理デバイスが接続されているポート23の番号が記録される。

【0020】

本実施形態では、1つの論理デバイスが2つのポート23に接続される、すなわち、1つの論理デバイスが2つのホスト計算機1からアクセスできる状態もとることができる。この場合、当該論理デバイスに対するテーブル情報が2つ存在する。図2に示す例では、論理デバイス番号2番のデバイスはポート番号0、及び1のポート23に接続されている。このため、論理デバイス番号2番の項目が2つ存在している。

【0021】

ターゲットID66とLUN67は、ホスト計算機1からアクセスする場合のSCSI-ID、LUNに相当する。1つの論理デバイスが複数のポート23からアクセスされる場合、それぞれのパス65に対応するターゲットID、LUNは図2に例が示されるように異なっていて良い。論理デバイス管理テーブルに保持された情報は、ネットワークインタフェース24を通じて、適当なタイミングで、あるいは記憶装置サブシステム2に障害が発生して構成が変化した、などの際に管理用ホスト計算機3へと送られる。このため、管理用ホスト計算機3も、論理デバイス管理テーブルを保持する。

【0022】

次に、管理用ホスト計算機3が、論理デバイス管理テーブル以外に保持する管理情報について説明する。

【0023】

管理用ホスト計算機3では、デバイスの割り当てを行うべきホスト計算機1ごとの情報を管理し、図3に示すようにホスト名71、ポート番号72、インタフェース番号73、論理デバイス番号74の組で構成されるホスト管理テーブルを持つ。

【0024】

ポート番号72と論理デバイス番号74は、いずれも記憶装置サブシステム2内で定義される番号で、これにより、各ホスト計算機1が、現在記憶装置サブシステム2の中のどの論理デバイスを割り当てられているか管理できる。インタフェース番号73は、各ホスト計算機1のインタフェース12を管理するために付けられた番号である。これは、1つのホスト計算機1が複数のインタフェース12を持つ場合があるため必要となる。ポート番号72とインタフェース番号73の組は、ホスト計算機1と論理デバイスの接続関係を示すために重要な要素である。例えば、図1におけるホスト計算機1bのように、2つのインタフェース12を持っていて、各インタフェース12から違うポート23を経由してある論理デバイスに接続されているという状況もありえる。このような構成を採ると、一方のインタフェース、あるいは、一方のインタフェースと記憶装置サブシステム2を接続するラインが障害などの理由で使えなくなった場合、他方のインタフェースから論理デバイスへの接続がなされていれば処理を続行でき、信頼性が高まる。

【0025】

ホスト管理テーブルと、記憶装置サブシステムから送られる論理デバイス管理テーブルを参照して、管理用ホスト計算機3は、各ホスト計算機1に論理デバイスを割り付けることができる。以下、デバイスの割り当ての処理について説明する。

【0026】

図4は、各ホスト計算機1a、1b、…で実施されるボリュームマネージャ11による処理の流れを示すフローチャートである。この処理は、ホスト計算機1を使用するユーザ、またはホスト計算機1で稼動するアプリケーションプログラ

ムなどが新規にデバイスを必要とするときに実施される。

【0027】

ステップ1001で、ボリュームマネージャ11は、ユーザ、またはアプリケーションプログラムから必要とするデバイスの数と各デバイスの種類の情報を得る。ユーザ、またはアプリケーションプログラムは、各デバイスの情報として、その容量、性能条件、信頼性レベルなどを指定する。

【0028】

容量は、先に説明した各デバイスのサイズのことである。性能条件としては、例えば、低速ディスク、高速ディスク、キャッシュ常駐ディスクを選ぶことができる。信頼性レベルとしては、例えば、RAID0、RAID1、RAID5、二重パス、リモートミラーを選ぶことができる。

【0029】

二重パスは、ホスト計算機が複数のインタフェースを持つ場合に選択でき、それを利用して同一デバイスに対し双方のインタフェースからアクセスできる。リモートミラーは、副記憶装置5に当該デバイスのコピーを持たせるものである。リモートミラーでは、地震、火災などの要因で記憶装置サブシステム2全体が稼働できなくなった場合にも、副記憶装置5にデータが保持されているため、信頼性を高めることが出来る。

【0030】

次に、ステップ1002でボリュームマネージャ11は、当該ホスト計算機1のインタフェース12上で使用されていないターゲットID、LUNの組を検索する。

【0031】

ステップ1003では、ステップ1001で指定された容量、性能条件、信頼性レベル、ステップ1002で検索された未使用のターゲットID、LUNの組を、管理用ホスト計算機3の管理マネージャ31に送信する。管理マネージャ31では、この情報をもとにして割り当てるべきデバイスが検索され、ホストインタフェース番号と、ターゲットID、LUNの情報を返却する。管理マネージャ31の処理については後述する。

【0032】

ステップ1004でボリュームマネージャ11は、管理マネージャ13からの情報を受け取る。そして、ステップ1005で、受け取った情報をもとに、デバイスが使用できるようにホスト計算機1の設定変更を行う。いわゆるオープン系のオペレーティングシステムなどの場合、ホスト計算機1が各デバイスにアクセスするために、デバイスファイルが各デバイス毎に用意され、このデバイスファイルに対してアクセスが実施される。これらデバイスファイルは、ホスト計算機1のデバイスコンフィギュレーション処理を行った際に用意される。これとは逆に、コンフィギュレーション処理時に存在しないデバイスに対しては、このファイルが存在しないために、ステップ1004でデバイスファイルの作成処理などを行う必要がある。例えば、サンマイクロシステムズ社のSolarisオペレーティングシステムでは、“drvconfig” コマンド、“drives” コマンドを発行することで、新規デバイスの認識、デバイスファイルの作成などが行われ、ホスト計算機1から新たに割り当てられたデバイスに対してアクセスできるようになる。

【0033】

最後にステップ1006で、割り当てられたデバイスファイル名、ターゲットID、LUNの情報をユーザまたはアプリケーションプログラムに通知してボリュームマネージャ11の処理が完了する。

【0034】

図5は、管理用ホスト計算機3上の管理マネージャ31による処理の流れを示すフローチャートである。

【0035】

管理マネージャ31は、ステップ1101でホスト計算機1から送られてきたデバイスサイズ、性能条件、信頼度レベルなどの情報を受け取る。ステップ1102では、管理マネージャ31内にある論理デバイス管理テーブル、及びホスト管理テーブルに設定されている情報をもとに、要求に応じたデバイスが存在するか検索する。検索対象となるデバイスは、「オフライン」のものである。そして、検索の結果、要求に合った「オフライン」状態のデバイスが見つかったかどうか判別する（ステップ1103）。

【0036】

要求に合った「オフライン」状態のデバイスが見つかった場合、管理マネージャ31は、ホスト計算機1から受け取った空きターゲットID、LUNの情報と、論理デバイス管理テーブル及びホスト管理テーブルに設定されている情報をもとにして、当該デバイスをホスト計算機1に接続するために適切なポート番号、ターゲットID、LUNを決定する（ステップ1104）。

【0037】

ステップ1105で管理マネージャ31は、当該論理デバイス番号のデバイスを、指定したポート番号、ターゲットID、LUNに設定し、オンライン状態にするよう記憶装置サブシステム2に指示する。記憶装置サブシステム2は、管理マネージャ31から受け取った情報に従って設定を行い、その結果を管理マネージャ31に返却する。

【0038】

ステップ1106で管理マネージャ31は、記憶装置サブシステム2から結果を受け取り、ステップ1107で、要求のあったホスト計算機1のボリュームマネージャ11に対して、インタフェース番号、ターゲットID、LUNを返却する。

【0039】

一方、ステップ1103で「オフライン」状態のデバイスで要求に合ったものが存在しなかった場合には、「未実装」の論理デバイス番号が存在するか検索する（ステップ1108）。「未実装」の論理デバイス番号が存在する場合には、記憶装置サブシステム2に対して、ホスト計算機1から要求のあったデバイスサイズ、性能条件、信頼度レベルなどの情報を伝えてデバイスの構築を指示する（ステップ1109）。

【0040】

記憶装置サブシステム2は、管理マネージャ31からの要求に合わせて当該デバイス番号のデバイスを構築し、管理マネージャ31に返却する。ステップ1110で管理マネージャ31は、その結果を受け取ると（ステップ1110）、ステップ1104の処理に移り、ポート番号、ターゲットID、LUNの設定を行

って、以下ステップ 1107 まで上述した処理を実行する。

【0041】

図 6、図 7 には、ホスト計算機 1 で不必要になったデバイスの返却処理のフローチャートである。

【0042】

図 6 は、デバイスの返却処理におけるボリュームマネージャ 11 の処理の流れを示すフローチャートである。

【0043】

デバイスの返却処理では、まず、ステップ 1201 でボリュームマネージャ 11 が、ユーザあるいは上位アプリケーションプログラムから不必要になったデバイスの情報、例えば、デバイスファイル名を受け取る。ボリュームマネージャ 11 は、ステップ 1202 でその情報をもとに、インタフェース番号、ターゲット ID、LUN を割り出す。

【0044】

ステップ 1203 でボリュームマネージャ 11 は、ホスト計算機 1 でそのデバイスを使わないようにするため、ホスト計算機 1 の設定変更を必要に応じて行う。この処理としては、例えば、デバイスファイルの削除などが行われる。

【0045】

ステップ 1204 でボリュームマネージャ 11 は、管理マネージャ 31 にインタフェース番号、ターゲット ID、LUN を通知し、処理を終了する。

【0046】

図 7 は、デバイスの返却処理における管理マネージャ 31 の処理の流れを示すフローチャートである。

【0047】

管理マネージャ 31 は、ステップ 1301 でホスト計算機 1 からインタフェース番号、ターゲット ID、LUN を受け取る。ステップ 1302 で管理マネージャ 31 は、受け取ったインタフェース番号、ターゲット ID、LUN をもとにして、記憶装置サブシステム 2 に対し、対象のデバイスをオフラインにするように指示する。そして、ステップ 1303 で記憶装置サブシステム 2 から、デバイス

をオフラインにした情報を論理デバイス管理テーブルの形式で受け取り、処理を完了する。

【0048】

以上、第1の実施形態を説明したが、管理マネージャの機能は必ずしも管理用ホスト計算機に存在する必要はなく、ホスト計算機1a、1b、…のいずれかに存在するように構成しても良い。また、図8に示すように管理マネージャを記憶装置サブシステムに設けるように構成しても構わない。この場合、各ホスト計算機1a、1b、…は、インタフェースを介して直接記憶装置サブシステムとの間で、要求の送出、情報受け取りを実施する。

【0049】

ー第2の実施形態ー

図9は、本発明の第2の実施形態における計算機システムの構成図である。ホン実施形態における計算機システムは、複数のホスト計算機、ホスト計算機1a、ホスト計算機1b、…、ホスト計算機1nと、複数の記憶装置サブシステム2a、2b、…、2m、管理用ホスト計算機3、ネットワーク4、ファイバチャネルスイッチ6を有して構成される。

【0050】

各ホスト計算機1a、1b、…は、第1の実施形態と同じく、ボリュームマネージャ11を有する。ボリュームマネージャ11は、管理用ホスト計算機3における管理マネージャ31と通信しあって動作する。さらに、ホスト計算機1は、インタフェース（図中、I/Fと略記）12を有し、インタフェース12によりファイバチャネルスイッチ8と接続される。

【0051】

記憶装置サブシステム2a、2b、…、2mは、それぞれ、第1の実施形態における記憶装置サブシステム2と同じく、ディスク21、ディスクコントローラ22、ポート23、ネットワークと接続するネットワークインタフェース（図中「ネットワークI/F」と略記）25から構成される。第1の実施形態と同じくディスク21、ポート23は複数あってもよいが、ここでは簡単のため各ディスク、ポートとも1つとして説明する。

【0052】

ファイバチャネルスイッチ 8 は、複数のポート 8 1 を有し、ホスト計算機 1 a、1 b、…のインタフェース 1 2、あるいは、記憶装置サブシステム 2 a、2 b、…と接続される。ファイバチャネルスイッチ 8 はまた、ネットワークインタフェース 8 2 を有し、ネットワークにも接続される。ファイバチャネルスイッチ 8 は、各ホスト計算機 1、各記憶装置サブシステム 2 が自由にアクセスできるようにするために使用される。この構成では、基本的にすべてのホスト計算機 1 が、すべての記憶装置サブシステム 2 に対してアクセスすることが可能である。

管理用ホスト計算機 3 は、第 1 の実施形態と同じく管理マネージャ 3 1 を有し、各ホスト計算機 1 a、1 b、…のボリュームマネージャ 1 1 と通信しあって動作する。

【0053】

図 10 は、管理用ホスト計算機 3 が保持する論理デバイス管理テーブルの一例を示すテーブル構成図である。この論理デバイス管理テーブルは、第 1 の実施形態において記憶装置サブシステム 2 が保持する論理デバイス管理テーブルに相当するテーブルである。以下、第 1 の実施形態における論理デバイス管理テーブルとの相違点について説明する。

【0054】

本実施形態では、管理用ホスト計算機 3 は、計算機システム内の全記憶装置サブシステム 2 の全デバイスに一意的な番号をつけて管理し、各デバイス毎にサイズ 1 0 3、構成 1 0 4、状態 1 0 5、LUN 1 0 6、WWN (World Wide Name) 1 0 2、及び接続ホスト名 1 0 7 を持つ。サイズ 1 0 3、構成 1 0 4、状態 1 0 5、LUN 1 0 6 については、第 1 の実施形態におけるものと同じものである。WWN 1 0 7 は、N_PORT_NAME と呼ばれ、各記憶装置サブシステム 2 のポート 2 3 に設定されている情報で、それぞれのファイバチャネルインタフェースで固有な情報である。接続ホスト名 1 0 7 は、当該デバイスに接続が許可されているホスト計算機のホスト名である。

【0055】

基本的に、ファイバチャネルスイッチ 8 で複数のホスト計算機 1、複数の記憶

装置サブシステム 2 が接続された構成では、各ホスト計算機 1 は、自由に各デバイスにアクセスできるが、システムの安全性に問題がある。このようなシステムの安全性に関する問題を解決するための手段として、例えば、特開平 10-333839 号公報には、ファイバチャネルで接続された記憶装置に対し、特定のホスト計算機からのアクセスだけを許可するような手段が開示されている。本実施形態の記憶装置サブシステム 2 も、システムの安全性を維持するため、特開平 10-333839 号公報に開示されているような手段を持っているものとするが、これについては、本発明の本質とは直接関係するところではなく、ここでは詳しい説明を省略する。

【0056】

さらに、WWN 107 は、各ホスト計算機 1 のインタフェース 12 にも与えられており、管理用ホスト計算機 3 は、図 11 に示すテーブルにより、ホスト名と WWN の組を管理している。

【0057】

以下、ボリュームマネージャ 11、管理マネージャ 31 の動作について説明する。

【0058】

本実施形態におけるボリュームマネージャ 11 の動作は、図 4 に示した第 1 の実施形態におけるものと基本的に同じである。ボリュームマネージャ 11 は、ユーザまたはアプリケーションプログラムから必要とするデバイスの数や種類の情報を受け、それをもとに管理マネージャ 31 に要求を送信する。管理マネージャ 31 の割り当てが終わると、デバイスの設定変更が行われて処理が終了する。

【0059】

管理マネージャ 31 も同様に図 5 に示した第 1 の実施形態における処理とほぼ同じ処理を行う。図 12 に、本実施形態における管理マネージャ 31 の処理のフローチャートを示す。なお、図 12 において、図 5 に示したものと同様の処理が行われる部分については、図 5 と同一の参照番号を用いている。以下の説明では、主に、図 5 に示す処理と異なる部分について説明し、図 5 と同一の処理が行われる部分については説明を省略する。

【0060】

管理マネージャ31は、ステップ1105で記憶装置サブシステム2にデバイスをオンラインにするよう指示する際、併せて記憶装置サブシステム2に当該ホスト計算機1からのアクセスを許可するように指示する。この指示では、管理マネージャ31から記憶装置サブシステム2に対して、アクセスを許可するホスト計算機1のWWNを通知する。各記憶装置サブシステム2は、不用意に他のホスト計算機からアクセスされないように、初期状態では全てのホスト計算機1からのアクセスを禁止している。記憶装置サブシステム2は、ホスト計算機1によるデバイスのアクセス時に、管理マネージャ31から受け取ったWWNに基づいて、そのアクセスの可否を判断する（ステップ2105）。

【0061】

さらに、管理マネージャ31は、ファイバチャネルスイッチ8の設定変更を行う。例えば、図13に示すように、ホスト計算機A、Bは、ディスクA、Bにアクセスするが、ホスト計算機Cは、ディスクCのみにアクセスする場合を考える。この場合、ホスト計算機Cに接続するポートcからは、ディスクA、Bに接続するポート（ポートd、ポートe）にはアクセスできないような経路設定を行って、あたかも2つのスイッチが存在するようにできる。これをゾーニングと呼び、これによってデバイスが本来アクセスが許されていないホスト計算機からアクセスされることを防ぐことができる。さらにデータの流れが分離されることにより、性能も向上する（ステップ2106）。

【0062】

異常の処理の後、管理マネージャ31の処理は、ステップ1106に移り、記憶装置サブシステム2からデバイス情報を受け取り、ターゲットID、LUN等の情報をボリュームマネージャ11に返却する。

【0063】

以上説明した実施の形態によれば、ホスト計算機がオンライン中にデバイスを要求にあわせて自由に割り当てることができる。さらにファイバチャネルスイッチで構成された、複数のデバイスが混在するような環境でも、要求に合ったデバイスを容易に割り当てることが可能となる。

【0064】

【発明の効果】

以上説明したように、本発明によれば、ホストコンピュータがオンライン中にデバイスを要求にあわせてに割り当てることができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施形態における計算機システムの構成例を示すブロック図である。

【図 2】

記憶装置サブシステム 2 により保持される論理デバイス管理テーブルの一例を示すテーブル構成図である。

【図 3】

管理マネージャ 3 1 により保持されるホスト管理テーブルの一例を示すテーブル構成図である。

【図 4】

ホスト計算機 1 内のボリュームマネージャ 1 1 による処理の流れを示すフローチャートである。

【図 5】

管理マネージャ 3 1 による処理の流れを示すフローチャートである。

【図 6】

デバイスの返却処理におけるボリュームマネージャ 1 1 の処理の流れを示すフローチャートである。

【図 7】

デバイスの返却処理における管理マネージャ 3 1 の処理の流れを示すフローチャートである。

【図 8】

本発明が適用された他の計算機システムの構成例を示すブロック図である。

【図 9】

本発明の第 2 の実施形態における計算機システムの構成例を示すブロック図で

ある。

【図 1 0】

管理マネージャ 3 1 により保持される論理デバイス管理テーブルの一例を示す
テーブル構成図である。

【図 1 1】

管理マネージャ 3 1 により保持されるホスト計算機 1 と WWN の対応関係を管
理するためのテーブルの一例を示すテーブル構成図である。

【図 1 2】

管理マネージャ 3 1 による処理の流れを示すフローチャートである。

【図 1 3】

ファイバチャネルスイッチ 8 のゾーニングの機能を示す説明図である。

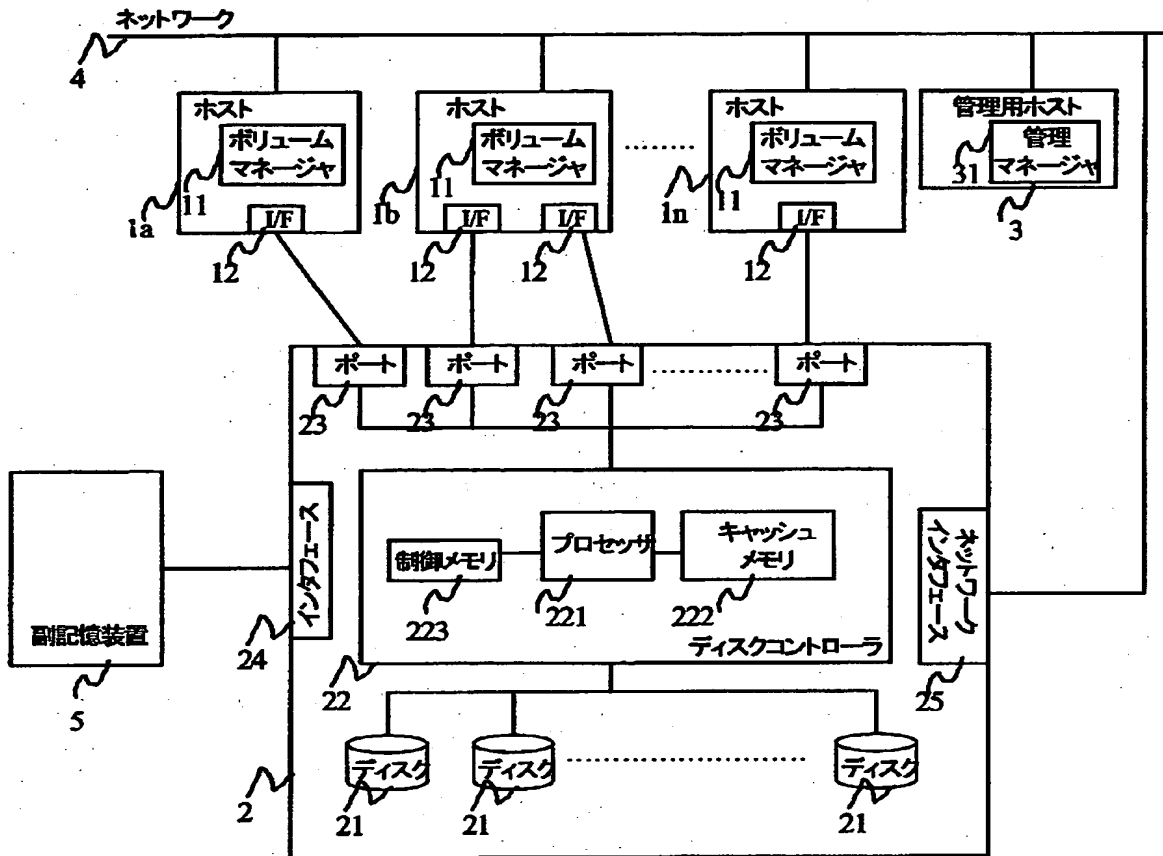
【符号の説明】

- 1 … ホスト計算機、
- 2 … 記憶装置サブシステム、
- 3 … 管理用ホスト計算機、
- 4 … ネットワーク、
- 5 … 副記憶装置、
- 8 … ファイバチャネルスイッチ、
- 1 1 … ボリュームマネージャ、
- 1 2 … インタフェース、
- 2 1 … ディスク、
- 2 2 … ディスクコントローラ、
- 2 3 … ポート、
- 2 4 … インタフェース、
- 2 5 … ネットワークインタフェース、
- 3 1 … 管理マネージャ、
- 8 1 … ポート、
- 8 2 … ネットワークインタフェース。

【書類名】 図面

【図 1】

図 1



【図 2】

図 2

| 論理デバイス番号 | サイズ | 構成 | 状態 | バス | Target ID | LUN |
|----------|---------|--------|---------|----|-----------|-----|
| 0 | 1000000 | RAID1 | オンライン | 0 | 0 | 0 |
| 1 | 1000000 | RAID5 | オフライン | - | - | - |
| 2 | 35000 | キャッシュ | オンライン | 0 | 0 | 2 |
| 2 | 35000 | キャッシュ | オンライン | 1 | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| k | - | - | 未実装 | - | - | - |
| k+1 | 1000000 | 単体ディスク | 障害オフライン | 0 | 1 | 0 |
| k+2 | 1000000 | RAID1 | オフライン | - | - | - |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| n | - | - | 未実装 | - | - | - |

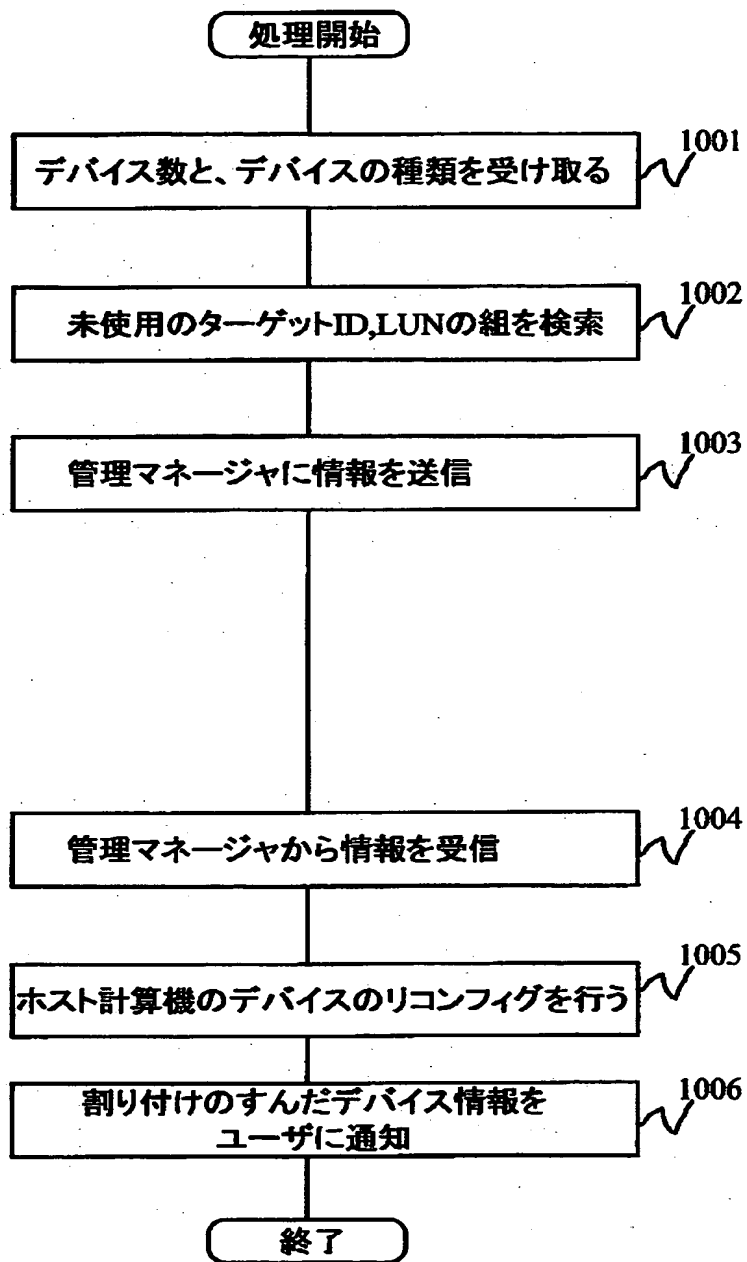
【図 3】

図 3

| ホスト名 | ポート番号 | インターフェース番号 | 論理デバイス番号 |
|-------|-------|------------|----------|
| host1 | 0 | 0 | 0,1 |
| host2 | 1 | 0 | 1 |
| host2 | 2 | 1 | 3 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| hostN | 0 | 0 | n |

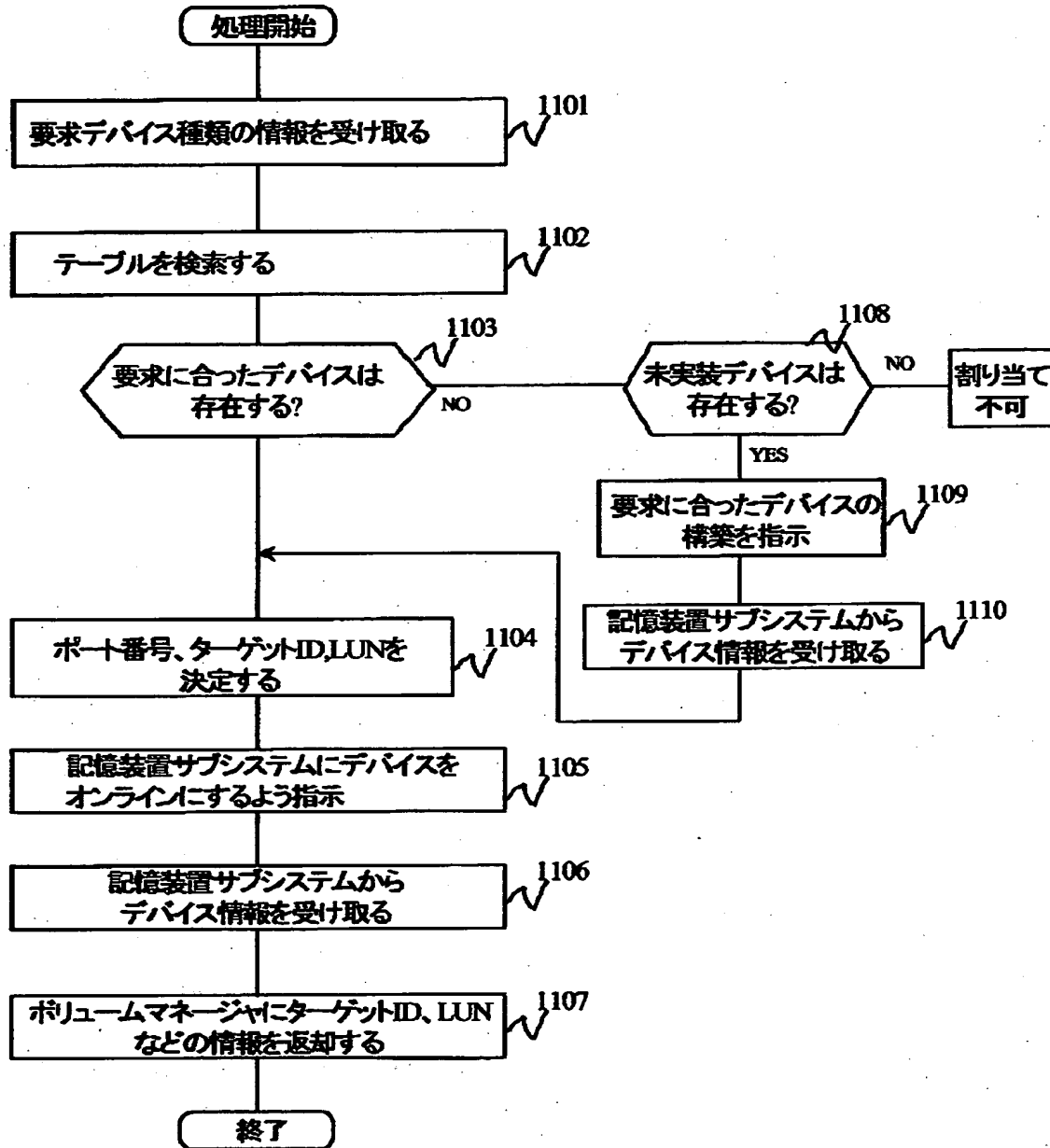
【図 4】

図 4



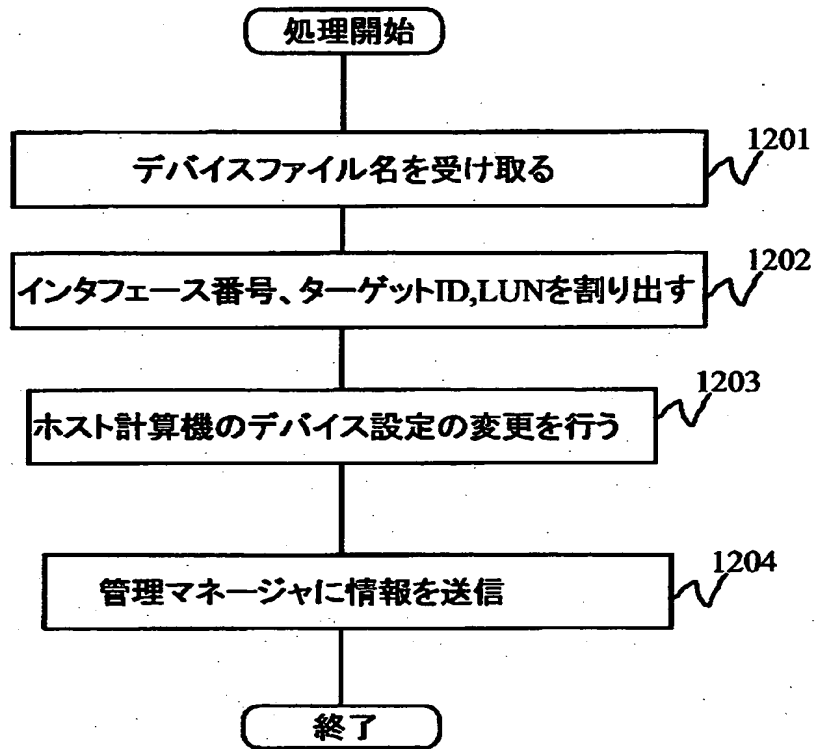
【図 5】

図 5



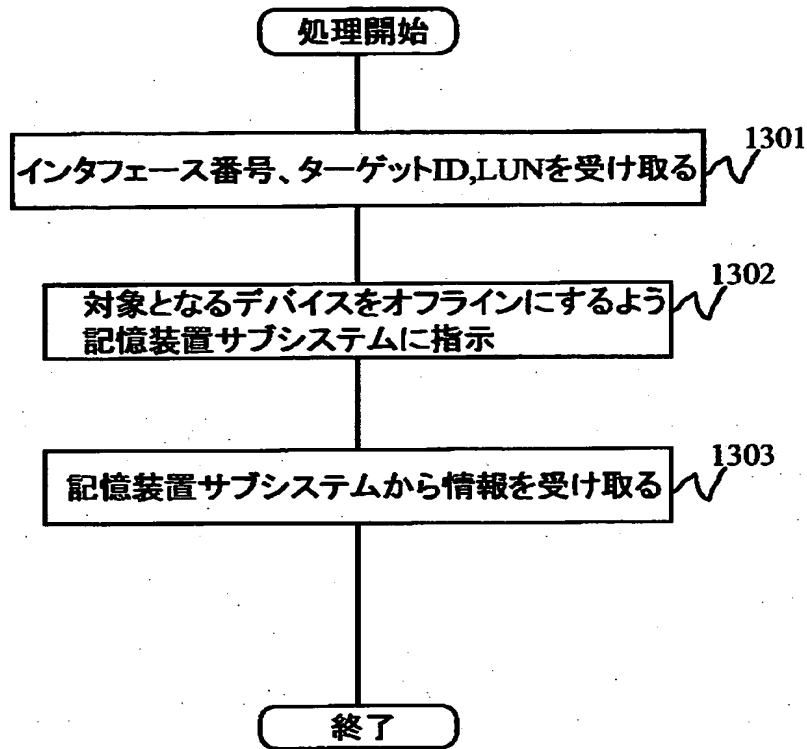
【図 6】

図 6

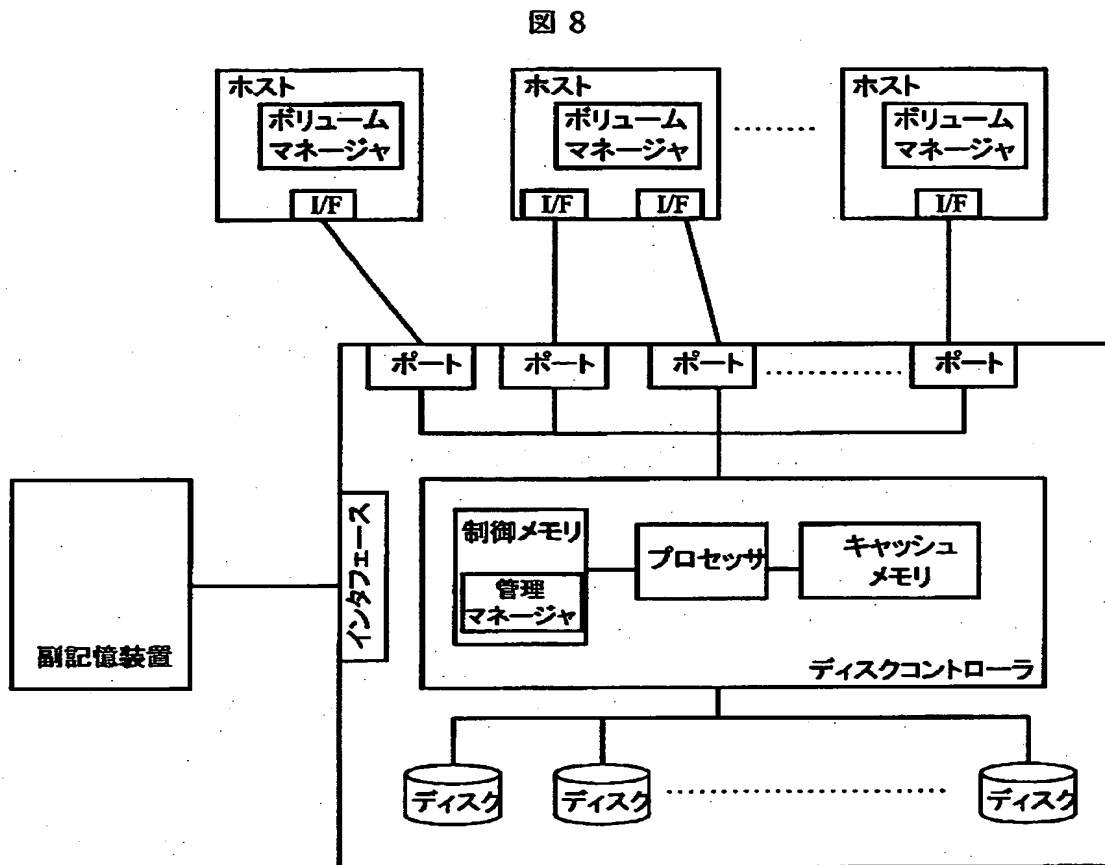


【図 7】

図 7



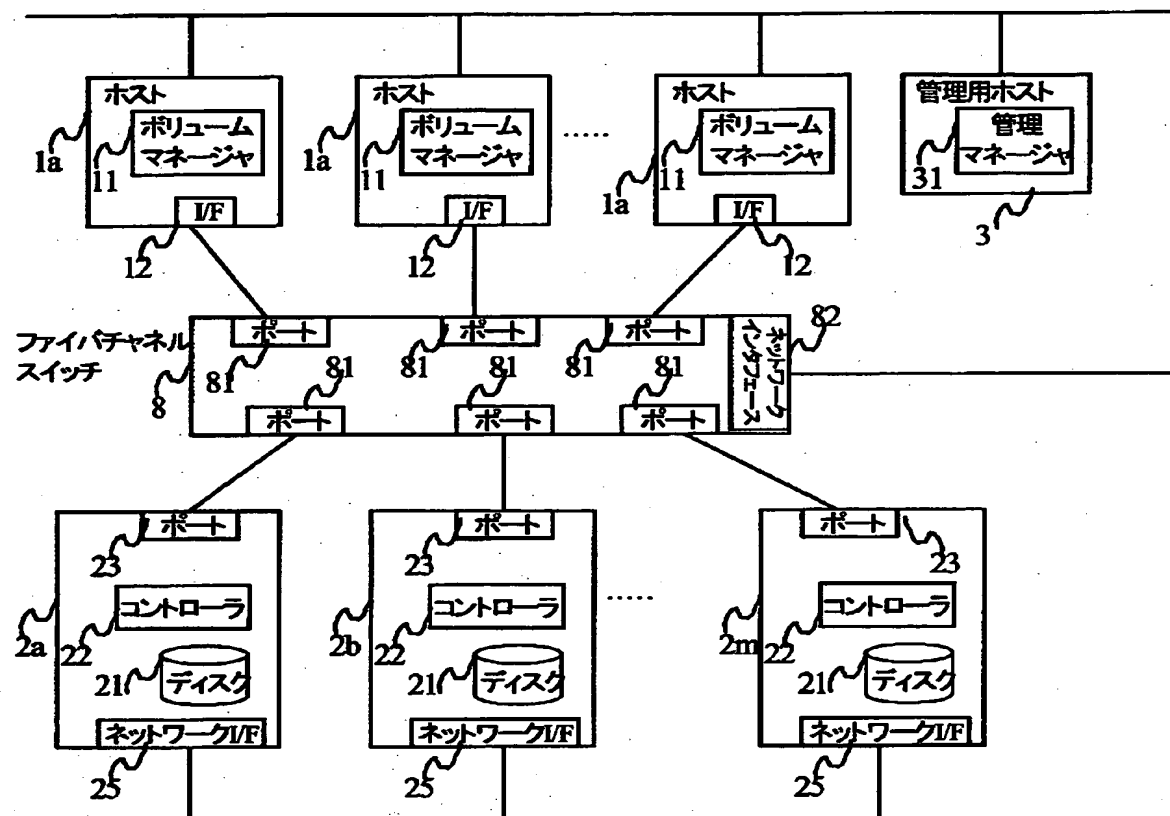
【図 8】



【図 9】

図 9

ネットワーク



【図 1 0】

図 1 0

| 101 デバイスNo. | 102 WWN | 103 サイズ | 104 構成 | 105 状態 | 106 LUN | 107 接続ホスト名 |
|----------------|------------|------------|-----------|-----------|------------|---------------|
| 0 | xxxxxxx | 1000000 | RAID1 | オンライン | 0 | hostA |
| 1 | xxxxxxx | 1000000 | RAID5 | オフライン | - | |
| 2 | xxxxxxx | 35000 | キャッシュ | オンライン | 2 | hostA,hostB |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| n | xxxxxxx | 1000000 | 単体ディスク | 障害オフライン | 0 | |

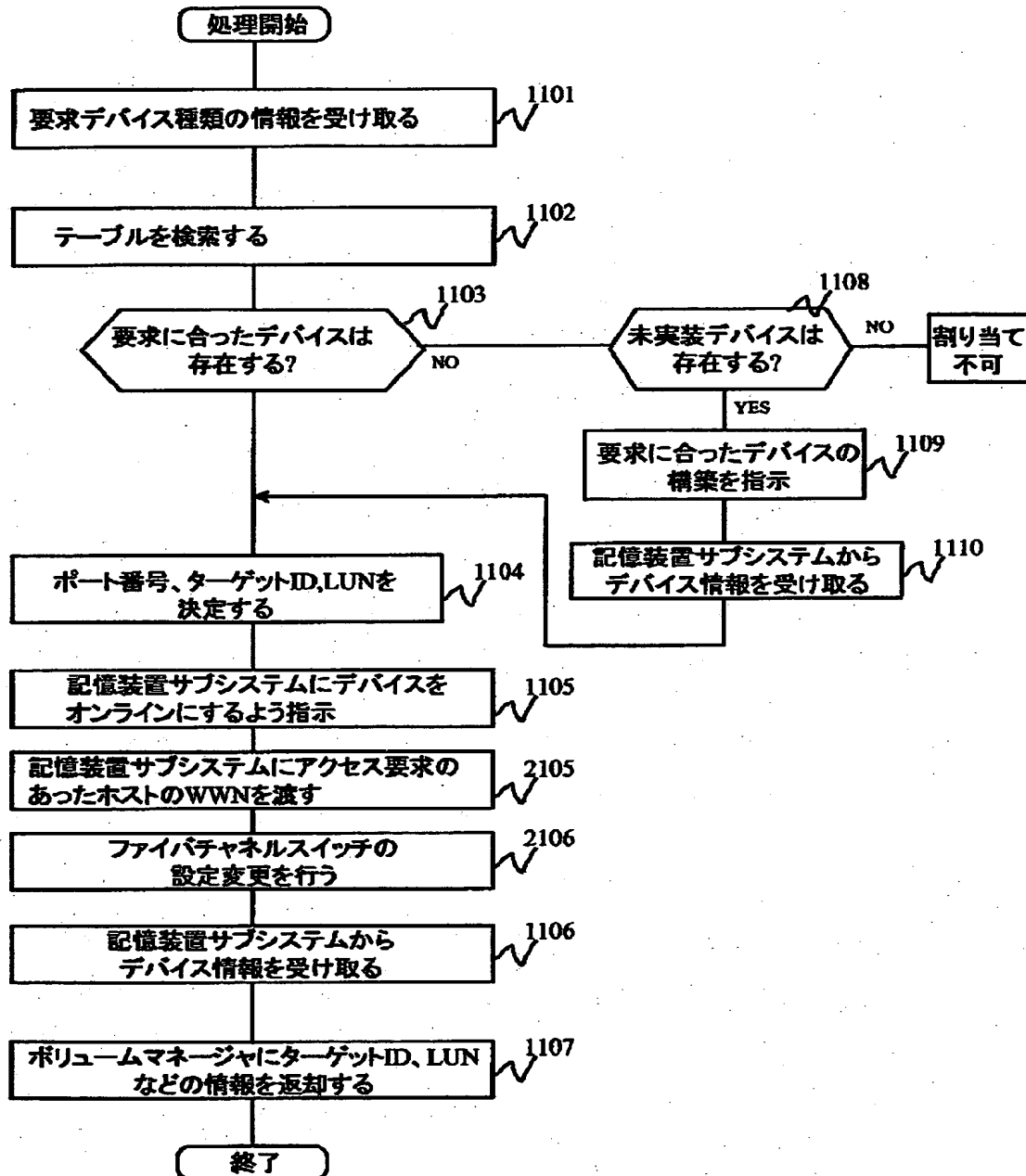
【図 1 1】

図 1 1

| 108 ホスト名 | 109 WWN |
|-------------|------------|
| host1 | xxxx |
| host2 | xxxx |
| host2 | xxxx |
| ⋮ | ⋮ |
| hostN | xxx |

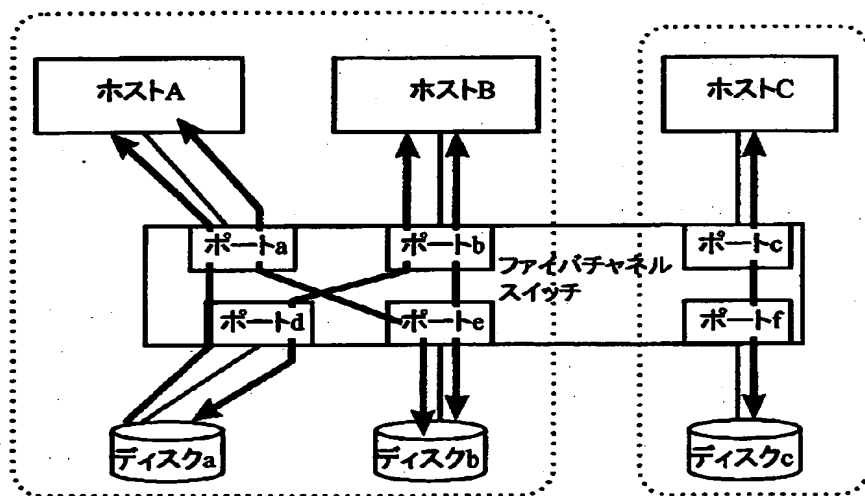
【図 12】

図 12



【図 13】

図 13



【書類名】 要約書

【要約】

【課題】

複数の計算機が用途に合ったデバイスを必要な時に必要な分だけ利用できるように、デバイスの設定、デバイスの各計算機への割り付けを動的に行えるようにする。

【解決手段】

記憶装置サブシステム 3 は種々の構成、容量を持つ複数のデバイスを有し、複数の計算機 1 に接続される。管理計算機 3 は、計算機 1 と記憶装置サブシステム 2 の接続関係を記憶・管理し、計算機 1 のボリュームマネージャ 1 1 と通信しながらデバイスの管理を行う。計算機 1 が新たにデバイスを要求する際、管理マネージャ 3 1 にデバイス数、容量、種類などを通知する。管理マネージャ 3 1 は、その要求に合ったデバイスを選択し、所定の情報を設定するよう指示する。その後、計算機 1 に所定の情報を返却し、ボリュームマネージャ 1 1 は、その情報に基づき計算機 1 の設定を変更しデバイスが使用できるようにする。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号

[000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所